

Екатерина Петкова
Институт за български език – БАН, София

ПРЕДСТАВЯНЕ НА ВЪТРЕШНОСИСТЕМНИТЕ ОТНОШЕНИЯ В ТЕРМИНОЛОГИЯТА В РАЗЛИЧНИ ВИДОВЕ ЕЛЕКТРОННИ ТЕРМИНОЛОГИЧНИ РЕСУРСИ

Abstract: The article examines the ways of presentation of the internal systematicity in terminology (relationships between concepts within a separate science domain), used in two different terminological knowledge bases – designed by corpus-based approach and by formalized approach. Few versions for presentation of the relationships between terms in formalized terminological knowledge base, derived from this two works, are proposed.

Keywords: terminological knowledge base, system approach

Бурното развитие на технологиите, което се наблюдава във втората половина на XX и в началото на XXI в., е свързано с постиженията на инженерните науки, получили от своя страна първоначален тласък от откритията на фундаменталните и приложните изследвания в различни области на естествените науки като физиката и математиката. Тяхното широко разпространение, наравно с развитието в социалната, политическата и икономическата сфера, допринася за ускоряване на процесите на глобализация, чието въздействие се отразява във всички области на обществения живот. Съвкупността от тези обстоятелства обуславя нарастващата необходимост от снабдяване на различни професионални и социални групи със специализирана терминологична информация.

Повод за настоящото изследване е необходимостта от дигитализиране на терминологичните архиви на Секцията по терминология и терминография към ИБЕ с цел те да бъдат предоставени за ползване в интернет от широки кръгове от обществото. В първата му част се разглеждат начините за представяне на вътрешната системност в терминологията, използвани в разработките на два екипа от учени, а именно в една корпуснобазирана и в една формализирана терминологична база знания (ТБЗ). Втората част предлага различни варианти за представяне на отношенията между термините, изработени въз основа на възприетото от тези разработки и въз основа на концепцията на М. Попова за речник от модулен тип (Попова 2006).

Най-напред трябва да се уточни значението на метатермините, използвани в заглавието на разработката. Под *терминология* тук се разбира съвкупността от термините, които означават понятията в отделна сфера на човешкото знание. Под *вътрешносистемни отношения* се имат предвид връзките между понятията в рамките на дадена област (в съпоставка

с външната системност, при която се отнася за връзките между понятия, принадлежащи към различни научни или професионални области). Към вътрешносистемните отношения спадат класификационните отношения, които са установени от човешкото мислене (родово-видовите отношения) и импликационните отношения, които се основават на реално съществуващи в действителността връзки между понятията (отношенията част – цяло и отношенията в предикатно-актантната сцена) (Попова 2012: 524 – 565). Под *терминологична база знания* (съкратено ТБЗ) се разбира електронно хранилище на знания, което съдържа богат набор от систематично структурирани терминологични данни, представящи понятията, техните характеристики и отношенията между тях по експлицитен начин (Мейер и др. 1997: 99). Под *формализиран подход* се има предвид такъв подход при изграждането на бази знания, чрез който организацията на понятията и отношенията между тях са представени систематично и във високо структуриран вид, напр. с използването на т.нар. концептуални графи (Мейер и др. 1997: 98; Кондамин, Ребейрол 1998: 29).

В настоящото изследване ще бъдат разгледани основните характеристики и особеностите на две различни разновидности терминологични бази знания, като първата от тях е изградена с използване на корпуснобазирания подход (Кондамин, Ребейрол 1998), а втората – с използване на формализиран подход (Мейер и др. 1997). Съпоставката им е направена при отчитането на следните критерии: 1) общи характеристики и подход при изграждането им; 2) цел, възможности за приложение и потенциални ползватели; 3) модел на разглежданата база знания (елементи и организация); 4) подход и метод на изграждане (вкл. етапи); 5) получени резултати; 6) възможности за допълване, модифициране и усъвършенстване.

Особености на терминологичната база знания, изградена въз основа на корпуснобазирания подход (Кондамин, Ребейрол 1998)

Общи характеристики. Тази ТБЗ не е формализирана, изграждането ѝ е независимо от дадено приложение. Представява съчетание между база данни и база знания. Консултациите със специалисти в разработваната област се извършват само в последния етап, за да се гарантира вярното отразяване на данните спрямо съдържанието на текстовете.

Цел, възможности за приложение и потенциални ползватели. Целта на разработката е да се създаде такъв модел ТБЗ, който позволява да се проследи как понятията функционират в даден текстов корпус. Тя е изпробвана и за създаване на индекс, с помощта на който да се правят справки за данните, съдържащи се в даден учебник или друг специализиран източник; за пренаписване на текст; за създаване на система за база знания за специално приложение (Кондамин, Ребейрол 1998: 29, 34 – 35).

Модел (елементи и организация на данните). Тази ТБЗ се състои от

четири модула. Първият модул съдържа *термините* (има се предвид само названието) – тук е поместена строго лингвистична информация, напр. род, вид, формални варианти, акроними, абривиатури и др. Вторият модул съдържа данни за *понятията*, които стоят зад термините. Данните се състоят от дефиниция и експлицитни понятийни-семантични отношения, наречени тук само концептуални отношения. Третият модул е озаглавен *връзка термини – понятия*, съдържа информация за степента на съответствие между термина (тук – названието) и описанието на конкретното понятие. Четвъртият модул е *корпусът* – съдържа текстове, които илюстрират по какъв начин термините се проявяват в отделни конкретни примери, т.е. показват как се реализира отношението „термин¹ – понятийно-семантично отношение – термин²“ (Кондамин, Ребейрол 1998: 30).

Подход и метод на изграждане (включително етапи). Използва се корпуснобазираният подход. Отделните етапи включват идентифициране на вероятните термини (с помощта на програмата LEXTER) и идентифициране на понятийно-семантичните отношения (с помощта на системата за анализ на текстове SATO) чрез задаване на определени модели на тези отношения.

Получените резултати се разпределят в три групи: 1) списък на термините; 2) списък на понятийните отношения с техните характерни модели – в това изследване присъстват родово-видовите, партитивните и отношенията между действие и негов вършител; 3) списък на двойките термини, свързани с дадено понятийно отношение (Кондамин, Ребейрол 1998: 32 – 34).

Възможности за допълване, модифициране и усъвършенстване. Според авторите на разработката тази терминологична база знания може да се допълва и модифицира в зависимост от нуждите на всеки отделен ползвател с цел достъп до знанията, съдържащи се в даден конкретен корпус (Кондамин, Ребейрол 1998: 34 – 35).

Особености на терминологичната база знания, изградена чрез формализиран подход с прилагане на системния анализ на областта (Мейер и др. 1997)

Общи характеристики. Тази ТБЗ е изградена със съвместните усилия на специалисти по терминология и по информационни технологии, които работят в посока формализиране на терминологичните данни. В работата по нея се използват т.нар. концептуални графи, представящи начина, по който понятията в дадена подобласт или в отделна нейна част са организирани в система. Поради прилагането на формализиран подход тази ТБЗ се определя като формализирана и ориентирана към приложенията. Системата за обработка на знания CODE (Conceptually Oriented Description Environment), използвана при създаването ѝ, осигурява средства както за съхранение и обработване на информация, така и за концептуален анализ на термините, който е от

най-съществено значение при изграждането на бази знания. Тази възможност се дължи на структурираната информация за признаците на понятията и отношенията между тях.

Цел, възможности за приложение и потенциални ползватели. Предназначена е на първо място да бъде в помощ на терминологите при анализ на системата от понятия в дадена област. Може да се използва от студенти при усвояване на дадена област на знанието чрез запознаване с нейните основни понятия и изучаване на системните отношения между тях, както и от преподаватели по терминология, по съставяне и редактиране на техническа документация и превод.

Модел (елементи и организация). Основни елементи на тази ТБЗ са понятията, характеристиките на понятията и отношенията между понятията. (По-подробно за формализирания подход към характеристиките и за видовете отношения между понятията, възприет от авторите на проекта, вж. Мейер и др. 1997: 100 – 104.) Видовете понятия автоматично „наследяват“ характеристиките на родовото понятие, като при въвеждане на характеристика и нейна стойност на едно ниво на родово-видовата йерархия тя се предава на останалите видови понятия (Мейер и др. 1997: 116).

Подход и метод на изграждане (включително етапи). Подходът при изграждането ѝ е йерархичен, основава се на резултатите от понятийния анализ на дадената специална област. Етапите, които включва, са: 1) Подготвяне на предварителен списък на термините, които ще бъдат застъпени, т.е. първоначално представяне на понятийната система на областта. Тази задача изисква усвояване на основните (ключовите) понятия и отношенията между тях, т.е. задълбочено изучаване на системата от понятия в разглежданата област. 2) Консултации със специалисти от разработваната научна област за откриване на несъответствия и неточности при представянето на понятийната система, за обсъждане на неясни елементи от нея. 3) Съставяне или подбор на дефиниции. В този етап познанията за понятийната система на областта са необходими при намирането на родовото понятие на описвания термин, на неговите еднородни (съподчинени на същото родово понятие) понятия и на характеристиките, които го разграничават от тези еднородни понятия. 4) Установяване на еквивалентите на термина в други езици чрез сравняване на характеристиките на понятието, означено с даден термин в единия език, с характеристиките на понятието, означено с неговия еквивалент в другия език. Ролята на системния понятийен анализ е от първостепенно значение при изпълнението на всяка от дейностите в изброените етапи (Мейер и др. 1997: 106 – 108).

Получени резултати. С помощта на тази система за обработка на знания могат да се правят *текстови и графични браузери*, които представят организацията на понятията в отделна научна област или в избрани нейни части, като възпроизвеждат обща картина на цялостната структура на знанията. Изградената чрез тази система терминологична база знания може да се редакти-

ра с цел добавяне на елементи, отстраняване на несъответствия в подредбата на понятията, в техните характеристики и взаимоотношения и др. (Мейер и др. 1997: 109).

Възможностите за допълване, модифициране и усъвършенстване на тази система за обработка на знания могат да се илюстрират с това, че при видовете понятия, които споделят повечето от характеристиките на родовото понятие, „наследяването“ на характеристиките става автоматично. При въвеждане на характеристика и нейна стойност на едно ниво на родово-видовата йерархия тя се предава на всички видови понятия, като може да бъде уточнена от терминолога на по-ниско ниво, а по този начин се дава възможност да бъдат регистрирани неточности и несъответствия в структурата на знанията (Мейер и др. 1997: 116).

При подробно запознаване с тези накратко представени разработки става ясно, че в резултат от използването на формализиран подход и поради подчертаното внимание към отразяването на системността в терминологията, създадената от Мейер и нейните сътрудници ТБЗ съдържа експлицирана йерархия на понятията в разглежданата специална област. Тук трябва отново да се подчертае, че това се дължи на системния понятиен анализ на дадената специална област, който е приложен при изпълнението на всяка от дейностите в изброените етапи.

По повод необходимостта от дигитализиране на терминологичните архиви на Секцията за българска терминология и терминография при Института за български език при консултации със специалисти по информационни технологии те поставят въпроса какви функции е нужно да поддържат планираните електронни ресурси и какво трябва да може да се извършва с тях, какви резултати трябва да се получават. В досегашните проучвания и тук се опитваме да дадем отговор на този въпрос.

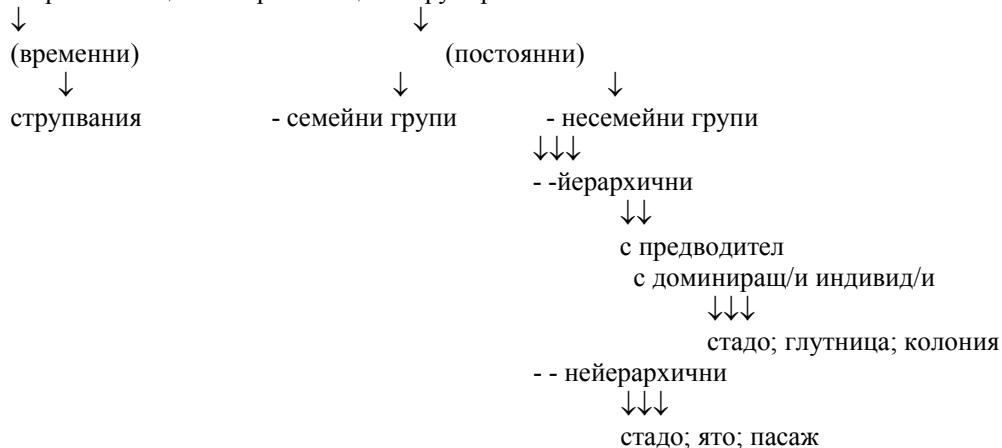
В резултат от проучванията на споменатите по-горе две и на още други разработки са изведени някои примерни начини за представяне на системността в терминологията, които да се използват в ТБЗ. Един от начините за сяга микроструктурата на отделната терминологична единица, показва какви елементи трябва да бъдат включени и как да бъдат структурирани (Петкова 2013). По-сложен е въпросът за макроструктурата на ТБЗ, в която трябва да се определи мястото на всеки от термините в понятийната йерархия на областта, т.е. да се отразят отношенията между тях (Петкова 2016). Предлаганите тук схеми са изработени въз основа на усвоеното от проучените досега различни ТБЗ и въз основа на системния подход към терминологията, създаден от М. Попова (Попова 2004; Попова 2006; Попова 2012: 524 – 565).

Едно от нещата, които са заимствани от разработката на Мейер (Мейер и др. 1997) в модифициран за нашите цели вид, е идеята за използване на различни видове браузери, с помощта на които (в зависимост от конкретния случай) се онагледяват различни особености на отношенията между понятията.

Напр. браузерите във вид на дървовидни схеми дават обща, генерална представа за подредбата на понятията в разглежданата област или нейна част.

Етологична структура на популациите

Форми на социална организация – групировки



За сравняване на характеристиките на понятия, подчинени на общ родов признак, може да се използват браузери в табличен вид. В таблицата по-долу са показани разликите в стойностите на отделните характеристики на понятията от най-ниското ниво в класификационната подгрупа „постоянни несемейни групи“, спадащи към *формите на социална организация* при животните.

Групировка	При кои видове	Присъствие на водач	Присъствие на доминиращ индивид	Социално разслоение	Морфологични различия
ято	птици	не	не	не	не
стадо	някои животни	да/не	да/не	да/не	не
пасаж	риби	не	не	не	не
глутница	някои животни	не	да	да	не
колония	някои животни	не	да	да	да

Въпреки невъзможността (поради ограничения обем) да се представи цялостна картина на отношенията между термините, която да показва всички равнища на понятийната йерархия в една плоскост, от приведените тук и в предишни разработки примери може да се съди, че терминологичните

бази знания са подходящи за отразяване на класификационната системност в терминологията. Същевременно на всяко ниво на разглеждане почти всички термини попадат в импликационни отношения с различни понятия както от други класификационни групи на дадената област, така и от други научни области. В следващи проучвания е необходимо да се намерят и подходящи начини за представяне на импликационните връзки между понятията. Опагледяването им вероятно не може да бъде направено по достатъчно подреден начин, подобен на отразяването на класификационните отношения, затова то следва да се осъществи с използването на семантични мрежи. Друга възможност е да се направи самостоятелен модул (подобно напр. на списъка на понятиятните отношения с техните характерни модели, вж. Кондамин, Ребейрол 1998: 32 – 34), в който да са изведени и илюстрирани с примери различни видове импликационни отношения между отделните понятия.

Според И. Мейер „всяка сериозна терминологична дейност изисква достигане на определена степен в усвояването на знания за дадената предметна област“ (Мейер и др. 1997: 98). Крайният резултат от терминологичните изследвания трябва също да бъде „кодирани като знание“, но „най-често по-голямата част от това знание, придобито и структурирано с усилен труд, остава там, където първо е била складирана, а именно в главата на терминолога. Само фрагменти от него са вложени в дефинициите или примерите.“ (пак там). Затова в заключение отново трябва да се подчертае необходимостта от развитие на теоретичните и приложните терминологични изследвания, които да задоволят нуждите на съвременното общество от специализирана терминологична информация, да го снабдят със средства за отразяване на знанията и да осигурят надеждна и ефективна професионална комуникация между специалистите от всички научни сфери.

СЪКРАЩЕНИЯ

ТБЗ – терминологична база знания

ЛИТЕРАТУРА

- Кондамин, Ребейрол 1998:** Condamines, A., Rebeyrolle, J. CTKB: A Corpus-based approach to a Terminological Knowledge Base. – In: *Computerm '98, First Workshop on Computational Terminology*. Montreal, Canada, 1998, pp. 29 – 35.
- Мейер и др. 1997:** Meyer, I., Eck, K., Skuce, D. Systematic Concept Analysis within a Knowledge-Based Approach to Terminology. – In: *Handbook of Terminology Management. Vol. I. Basic Concepts of Terminology Management*. 1997, pp. 98 – 118.
- Петкова 2013:** Петкова, Е. Видове електронни терминологични ресурси и възможности за тяхната употреба. – В: *70 години българска академична лексикография*. София: АИ „Проф. Марин Дринов“, 2013, с. 329 – 334.

- Петкова 2016:** Петкова, Е. Отразяване на понятийно-семантичната системност в терминологията в макроструктурата на формализирана терминологична база знания. – В: *За словото – нови търсения и подходи*. София: Изд. на БАН „Проф. М. Дринов“, 2015, с. 222 – 230.
- Попова 2004:** Popova, M. A System Approach in Terminography. – In: *Terminology at the Time of Globalization*. Ljubljana, 2004, pp. 33 – 39.
- Попова 2006:** Попова, М. Терминологичен речник от модулен тип. – В: *Светът на речника. Светът в речника*. Велико Търново: ИК „Знак’94“, 2006, с. 106 – 112.
- Попова 2012:** Попова, М. *Теория на терминологията*. Велико Търново: ИК „Знак’94“, 2012.